Joint Essential Tasks and A Framework for Evaluation

Gina Kingston and **Kevin Johns**

Joint Systems Branch
DSTO Fern Hill
Department of Defence
Canberra ACT 2600
Australia
Gina.Kingston@dsto.defence.gov.au

Strategic and Capability Studies Branch
DSTO Pyrmont
Locked Bag 12
PO Box 44 Pyrmont NSW 2009
Australia
Kevin.Johns@dsto.defence.gov.au

Abstract

In Jan 2000, DSTO embarked on a task to develop a list of Australian Joint Essential Tasks (AS JETs) to support the planning and evaluation of exercises. This process has been very successful and has seen ASJETs incorporated into the planning for the next major Australian-led exercise – Crocodile 03.

Evaluation may be used for a variety of purposes, for example: to monitor progress; to determine whether or not objectives can be achieved; to assess the suitability of new processes, systems or structures; and to identify deficiencies or dependencies. Each of these purposes has different evaluation requirements. Also evaluation may be limited to determining which activities have been conducted, or be based on subjective opinions, repeatable subjective measures, or objective measures.

While rigorous objective evaluation is often seen as the most desirable and reliable, it may prove impractical or expensive, so a risk-managed approach is required. The approach needs to take into account a number of factors including the purpose of the evaluation and the time and results of the last evaluation to determine which measures to use. With experience, it should be possible to identify a suite of reliable measures – both subjective and objective. This approach is discussed in the context of planned evaluations for future military exercises.

1. Introduction

Australian Joint Essential Tasks (ASJETs) have recently been developed to support the evaluation of exercises and operations in the Australian Defence Force (ADF). ASJETs were designed to capture those tasks that are essential to the preparation for, planning for, and conduct of, military operations.

ASJETs were developed from considerations of basic military tasks undertaken internationally and revised in workshops attended by military personnel from a variety of backgrounds. The ASJETs are collected into similar functional areas – Command and Control (C2), Intelligence, Shape, Conduct, Protect and Sustain at the Strategic, Operational and Tactical levels as shown in Figure 1. For more detail on the development of ASJETs see [Kingston *et al.*, 2002].

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
SEP 2002		**00-00-2002 to 00-00-2002**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Joint Essential Tasks and A Framework for Evaluation	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **DSTO Fern Hill,Department of Defence,Canberra ACT 2600 Australia, ,**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
7th International Command and Research Control and Technology Symposium

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**10**	

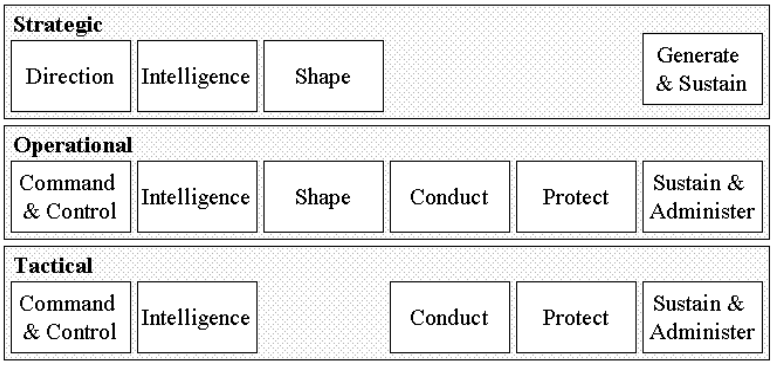

Figure 1. AS JETS Structure

This paper discusses the use of ASJETs (Section 2) and associated measures (Section 3) to support evaluation in military operations and exercises. Sections 4 and 5 discuss the current status and future directions for the evaluation of ASJETs.

2. **Background**

2.1 ***Evaluation in the ADF***

A stylized representation of the process envisaged is given in Figure 2.[1] During the planning of military exercises the ASJETs to be evaluated are carefully selected, the raw information to be obtained identified. The results are subsequently analyzed and metrics determined. The final stage, net assessment, is the consideration or consolidation of all of the measures associated with an ASJET to determine whether or not the ASJET was performed satisfactorily. Section 3 discusses the first steps towards this.

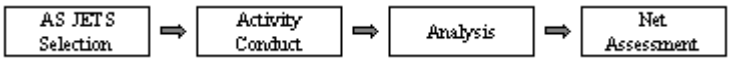

Figure 2. AS JETS driven evaluation.

2.2 ***Military Exercises***

The Australian exercise program runs on a four-year cycle, which includes an international exercise and two Australian-led exercises in which either the US or New Zealand participates. The planning, preparation, and seeding of information for the Crocodile 03 exercise have already commenced, and, while the field exercise is to be held in 2003, the Strategic planning phase of the exercise will be held this year.

These major exercises offer a unique opportunity to study Australia's military capability. They are of fixed duration, in a known location, and use a scenario deliberately chosen to exercise extant military capability. These factors mean that a more comprehensive evaluation can be planned and conducted than would be possible in smaller exercises or actual operations. It must, however, be

[1] In addition to being used for evaluation, AS JETs are influencing the development of doctrine.

recognised that they also fulfil a training role and may use highly stylized or artificial scenarios. A study of such exercises alone is thus not sufficient to assess military capability.

3. **Measurement**

This section introduces our approach to measurement, discusses the properties of good measures, and illustrates our approach to developing measures, using an ASJET that can then be evaluated in military exercises.

3.1 ***Measurement Hierarchy***

In order to conduct Net Assessment, a method of aggregating measures is required. The approach proposed in this paper is to utilise a generic set of performance criteria that can be readily understood. These performance criteria are then customised for each ASJET according to its functional area, and according to a process model view of an ASJET.

As shown in Figure 3, each ASJET can be thought of as a Process to achieve particular Outputs or Outcomes from given Inputs. While the ASJET measures should focus on the outputs and outcomes of a task, and measures for communications and electronics aspects of information often focus on parameters such as bandwidth, transmission rates, etc. [Ballenger, 2001], the consideration of Inputs can be important for Net Assessment. In particular, insufficient, or inappropriate Inputs may underlie poor performance. Consideration must also be given to the conditions under which the task was conducted, including the availability of Resources.

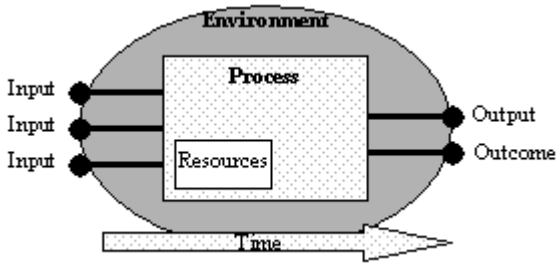

Figure 3. AS JETS viewed as a process model.

The process model as shown in Figure 3, also provides some insights into the performance criteria. The timeline suggests that the Inputs, Process and Outputs should be Timely. The solid circles on the connecting lines suggest that the Inputs and the Outputs should be Appropriate and Complete. Although, what it means for an Input, Output or even a Process to be Appropriate needs to be further elaborated.

Clark and Moon [Clark and Moon, 2002] have suggested that information, either as an Input or an Output needs to be Timely, Relevant, Accurate and Complete and have used these constructs to classify a range of measures. Similarly, other Inputs and Outputs should also be Relevant and Accurate.

	Input	Process	Output	Outcome
Command and Control	Relevance Accuracy	Integration Conformance Economy	Relevance Accuracy	Robustness Control Economy
Intelligence				Robustness
Shape				Stability Support
Conduct				Lethality Stability
Protect				Integrity
Sustain				Maintainability
SUMMARY	TRAC	TICEC	TRAC	O

Figure 4. Appropriateness Criteria by Functional Area and Process Model Perspective.

Processes should be Integrated both internally between organisations and externally with other processes; Conform to relevant guidance, plans, and doctrine; and be Economic with the use of resources. Outcomes, however, should be assessed in the context of the mission and the Functional Area. (See Figure 4.)

Command and Control (C2) is about the generation of plans and interaction with superiors, subordinates and other organisations. Accordingly plans should be relevant, pass information, be clear, be timely, be flexible and be economic. The output constructs, TRAC, capture most of these – with Accuracy the clarity of the information. The additional attributes of Robustness (flexibility), Control (passing of information) and Economy are associated with the outcomes of the C2 tasks. While Plans should also enable the success of an operation, they cannot ensure it, hence 'Plans' has not been included in Figure 4.

Intelligence comprises direction, collection, processing and dissemination of information to users and decision-makers [ADFP 101, 1994]. It must be timely, reliable and accurate.[2] Since the purpose of intelligence is to reduce uncertainty and risk in military operations, it should also be Robust to changing assumptions, inaccurate or imprecise information and changing requirements.

Shaping the environment is about the creation of a strategic environment that supports Australia's long-term security interests. Shaping thus aims to build support among nations for common security issues and reduce the potential for misunderstandings and the escalation of tension. To this end, Stability and Support are the hallmarks of successful Shaping activities. Conducting military exercises may be about obtaining Stability (for example in Peace Keeping operations) or about Lethality. Lethality, Power and Success must also be considered in addition to Timeliness (Time and Rate) and Completeness. For countries with smaller forces, such as Australia, the focus is on appropriate and measured response to obtain desired effects rather than the application of overwhelming power. As previously discussed, we have deliberately excluded Success from our model. 'Protect' is about maintaining the integrity of our infrastructure, vital assets and personnel.

For 'Sustain' tasks the outputs should be readiness, responsiveness and the sustainability of logistics support. Readiness and responsiveness are both aspects of Timeliness, so the additional

[2] Reliability is covered under Accuracy in our framework.

outcome of Maintainability (i.e., an ability to maintain and sustain operations over a long period of time) has been included.

AS JETs	AS JET A			AS JET B			AS JET C			
	Input	Process	Output	Input	Process	Output	Input	Process	Output	
Constructs	TRAC	TICEC	TRACO	TRAC	TICEC	TRACO	TRAC	TICEC	TRACO	…
Measures	M1.I1	M1.P1	M1.P1	M2.I1	M2.P1	M2.P1	M3.I1	M3.P1	M3.P1	
	M1.I2	M1.P2	M1.P2	M2.I2	M2.P2	M2.P2	M3.I2	M3.P2	M3.P2	
	M1.I3	M1.P3	M1.P3	M2.I3	M2.P3	M2.P3	M3.I3	M3.P3	M3.P3	
	M1.I4	M1.P4	M1.P4	M2.I4	M2.P4	M2.P4	M3.I4	M3.P4	M3.P4	
	M1.I5	M1.P5	M1.P5	M2.I5	M2.P5	M2.P5	M3.I5	M3.P5	M3.P5	
	M1.I6	M1.P6	M1.P6	M2.I6	M2.P6	M2.P6	M3.I6	M3.P6	M3.P6	
	…	…	…	…	…	…	…	…	…	

Figure 5.The relationship between AS JETs, Constructs and the Measures.

For each of the constructs described above, measures need to be defined that capture the required attributes of the constructs. Measures and their associated measurement techniques, will generally be specific to an ASJET. Other measures, related to the goals of an exercise, operation, or experiment, may also be included. Figure 5 gives a summary of the relationships between AS JETs, the constructs described above and measures.

As discussed, the proposed method is based on a transition from performance criteria, expressed in lay terms, to measures of performance. This method can be mapped to other approaches, such as the Military Operations Research Society (MORS) hierarchy of measures as discussed by Clark and Moon [Clark and Moon, 2002].

3.2 *What makes a good measure?*

A good measure should accurately capture the property of interest. This means that the measure should correspond with our intuitive understanding of the property, and that it is possible to compare values for the metric obtained by different people, at different times. More formally, it should have face (or content) validity and be repeatable. The measure should also have predicative and construct validity, that is, it should obey known relationships between the property being measured and other properties [Ellis, 1994; Straub, 1989].

It is often assumed that objective measures – where an independent measuring instrument is used[3] – are better than subjective measures, for which no measuring instrument exists. However, subjective measures have a role, as they often align with out intuitive understanding of a property and therefore have *prima facie* credibility and can be readily understood. It should be noted that objective measures are also prone to a range of errors as shown in Figure 6. For example, two students reading the dial on an analogue Ohmmeter may read a different value if one reads the dial from an angle and the other from face-on (called parallax error). If the device was not correctly set to zero, then both readings will be incorrect owing to an 'Instrumentation Error'.

[3] Such as a rule or a thermometer.

Self Assessment	Subjective	Objective
• Observational Error o Failure to understand the question. o Failure to recall an answer. o Indecision. • Expectation Error • Instrumentation Error o Dishonesty. [Ellis, 1994]	• Observational Error • Expectation Error • Instrumentation Errors (Attitude) o Biosocial o Situational [Christensen, 1980]	• Observation Error o Failure to read the results correctly. • Instrumentation Error (measurement instrument gives an inaccurate reading) [Ellis, 1994]

Figure 6. Standard measurement errors.

Subjective measures may be repeatable, particularly when they are based on well-understood principles or when the inputs from multiple sources are obtained. Few measures are, however, entirely subjective or objective.

Consider the decision on which colour to choose when buying a car so as to obtain the best resale value. You could make a subjective decision, based on which colour you think looks the best, conduct a survey of which colour people generally prefer or you could make a decision based on an objective measure such as the number of cars sold in each colour. While the last measure is objective, it could be less reliable than the more subjective measures owing to unknown factors such as the number of fleet cars included in the sales data.

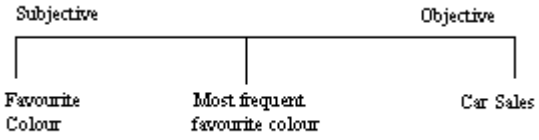

Figure 7. Measures of car colour for resale value.

Assuming that we could obtain values for all three metrics, we could determine how well they correlate. If all three measures consistently gave the same value (for example, in different countries), we could use any of the measures for forecasting car resale value. If they produced different values, it is necessary to determine why. For example, if car sales included lots of white fleet cars with low resale value, it might be best to use the most frequent favourite colour. If it was also true that a random sample could not be obtained to assess the most frequent favourite colour, then it might be best to add the rankings of the colours using both these measures, and to choose the colour with the lowest total score. If we had historical information on which approach was best, we could make some decisions about the cost and benefit of each of the measures and select the most cost effective measures.

3.3 *Selecting Suitable Measures - An Example*

To evaluate the Immediate Planning Process in a military exercise consider the constructs: ‘Process Completeness’ and ‘Outcome Economy’ in more detail.

3.3.1 *Process Completeness*

Process Completeness complements the constructs of Input Completeness, Output Completeness and Process Conformance. Input Completeness considers the availability and use of resources and Output Completeness considers the availability and use of resources for other tasks, so Process Completeness focuses only on the *activities* undertaken, and then only on the relationships between activities planned, commenced and completed. Process Conformance considers whether or not the process followed relevant guidance, doctrine and procedures.

Process completeness is captured by a series of groups of measures:

- The percentage of activities planned but not commenced.
- The percentage of activities planned but not completed.
- The percentage of activities commenced but not completed.

Activities conducted, but not planned, are covered under Process Conformance.

In some circumstances, it may be possible to further elaborate these to consider either the degree to which activities are completed, or critical activities.

- The percentage of activities planned but not commenced.
    - The percentage of critical activities planned but not commenced.
    - The percentage of non-critical activities planned but not commenced.
- The percentage of activities planned but not completed.
    - The percentage of critical activities planned but not completed.
    - The percentage of non-critical activities planned but not completed.
    - The degree to which planned activities are completed.
    - The degree to which planned critical activities are completed.
    - The degree to which planned non-critical activities are completed.
- The percentage of activities commenced but not completed.
    - The percentage of critical activities commenced but not completed.
    - The percentage of non-critical activities commenced but not completed.
    - The degree to which commenced activities are completed.
    - The degree to which commenced critical activities are completed.
    - The degree to which commenced non-critical activities are completed.

Amongst other things, the measurement approach needs to consider how to determine what activities are planned, how to distinguish between activities, how to determine when they are complete, and possibly how to determine the degree of completeness and how to identify critical activities. Here we discuss measurement approaches for the percentage of activities planned but not commenced.

As shown in Figure 8, the planning process may be divided into seven activities. Here it is assumed that all of these need to be conducted.

1. Determine enemy Centres of Gravity
2. Determine friendly forces Centres of Gravity
3. Determine the outcome sought
4. Identify Doctrinal Options
5. Develop Courses of Action
6. Produce Military Strategic Guidance
7. Determine military forces to be assigned

Figure 8. Strategic-level Immediate Planning Process

We now discuss methods for determining whether or not these tasks are complete.

Method 1. Look at the number of hours spent on the activities, compared to the usual duration. The approach has the benefit that it is simple to measure. However, it requires background information. Furthermore, the method is of limited use, unless the cause for deviations can be determined. For example, what does it mean if an activity takes longer than normal?

Method 2. Use independent expert opinion. This approach is also simple to measure and is more robust to than the previous method, provided that suitable experts can be found. It is subjective, but may be repeatable if multiple experts are used. Furthermore, if the experts can justify their opinions, it may provide the basis for determining more objective measures in the future.

Method 3. Consider the availability of activity outputs. This approach considers only the availability and not the quality of the outputs, which are Output constructs. This approach is objective, but in only suitable for activities that produce distinct outputs.

Method 4. Consider the commencement or completion of dependent tasks or steps. This approach is also objective, but while it does not require concrete outputs, it does require an understanding of work-flow procedures. For example, Activity 6 in Figure 8 can be commenced, but not completed before the completion of the preceding steps.

For evaluating military exercises we suggest using a combination of Approaches 2, 3 and 4. In ongoing work we plan to look for consistency between these three approaches, and try to identify when each approach is most useful.

3.3.2 *Outcome Economy*

While Process Completeness was closely related to several other constructs, ‘Output Economy’ is relatively independent from the other constructs. It is concerned with the economic use of resources, or in this case with the indicative resources required to fulfil the Military Strategic Planning Guidance (MSPG).

Output Economy is captured by a series of groups of measures:

- The cost of alternative courses of action as a percentage of the cost of the course of action recommended in the MSPG.

- The resources indicated for alternative courses of action as a percentage of the resources indicated by the MSPG.
- The cost of repairing or replacing resources that are damaged, destroyed, or are otherwise unable to be used for subsequent courses of action as a percentage of that indicated by the MSPG.

Note that these measures have been designed so that higher values are more desirable, for consistency with Process Completeness. However, whereas a maximum value of 100% was possible for Process Completeness, it was decided that it was more important to keep the measures simple rather than to manipulate the measures to retain a maximum value of 100%.

Once again, expert opinion could be used to assess these measures. There are also a number of more robust approaches that could be used. For this AS JET and construct, it may, however, be necessary to rely on expert opinion to generate alternative courses of action.

Method 5. Detailed economic indicators for the course of action recommended in the MSPG and alternative courses of action could be assessed through a number of methods including war gaming, simulation, or the conduct of multiple exercises or experiments.

For future military exercises will we use both expert opinion and the results of the war-gaming conducted during the exercise. It is considered that this approach can adequately assess the cost-effectiveness of the immediate planning process, where decisions need to be made relatively quickly. There are, however, potential benefits to conducting more detailed evaluations at a later date or if the results were to be used for contingency planning.

4. Future Directions

We plan to assess this approach in future military exercises and to estimate the cost of both deriving measures according to this framework for all ASJETs, and the cost of evaluating them in a major exercise. The results from the evaluation will be fed into the ADF's exercise repositories, and it is hoped that in the future, these systems will provide a repository of measures, their values, their validity, and the effort required to collect and evaluate them. Future exercise planners can then select the most appropriate measures to meet the exercise evaluation objectives.

Acknowledgement The authors wish to thank the helpful suggestions from a referee.

5. References

[ADFP 101, 1994] Australian Defence Force Publication (ADFP) 101 – Glossary, May 1994.

[Ballenger, 2001] J.P. Ballenger. 'Reconnaissance by Fire: What Does the Future Hold?'. *Phalanx*, p. 4, March 2001.

[Christensen, 1980] Larry B Christensen. *Experimental Methodology*. Allyn and Bacon Inc. 1980.

[Clark and Moon, 2002] Thea Clark and Terry Moon. *Assessing the Military Worth of C4ISR Information*. International Command and Control Research and Technology Symposium, Quebec City, Canada, 2002.

[Ellis, 1994] Lee Ellis. *Research methods in the Social Sciences*. WCB Brown & Benchmark. 1994.

[Kingston, McCarthy and Johns, 2002] Gina Kingston, Anna McCarthy and Kevin Johns. *Developing and Evolving and Joint Task List*. International Command and Control Research and Technology Symposium, Quebec City, Canada, 2002.

[Straub, 1989] Detmar W Straub. *Validating Research Instruments*. MIS Quarterly, Vol 13, Issue 2, pp 147-169. 1989.